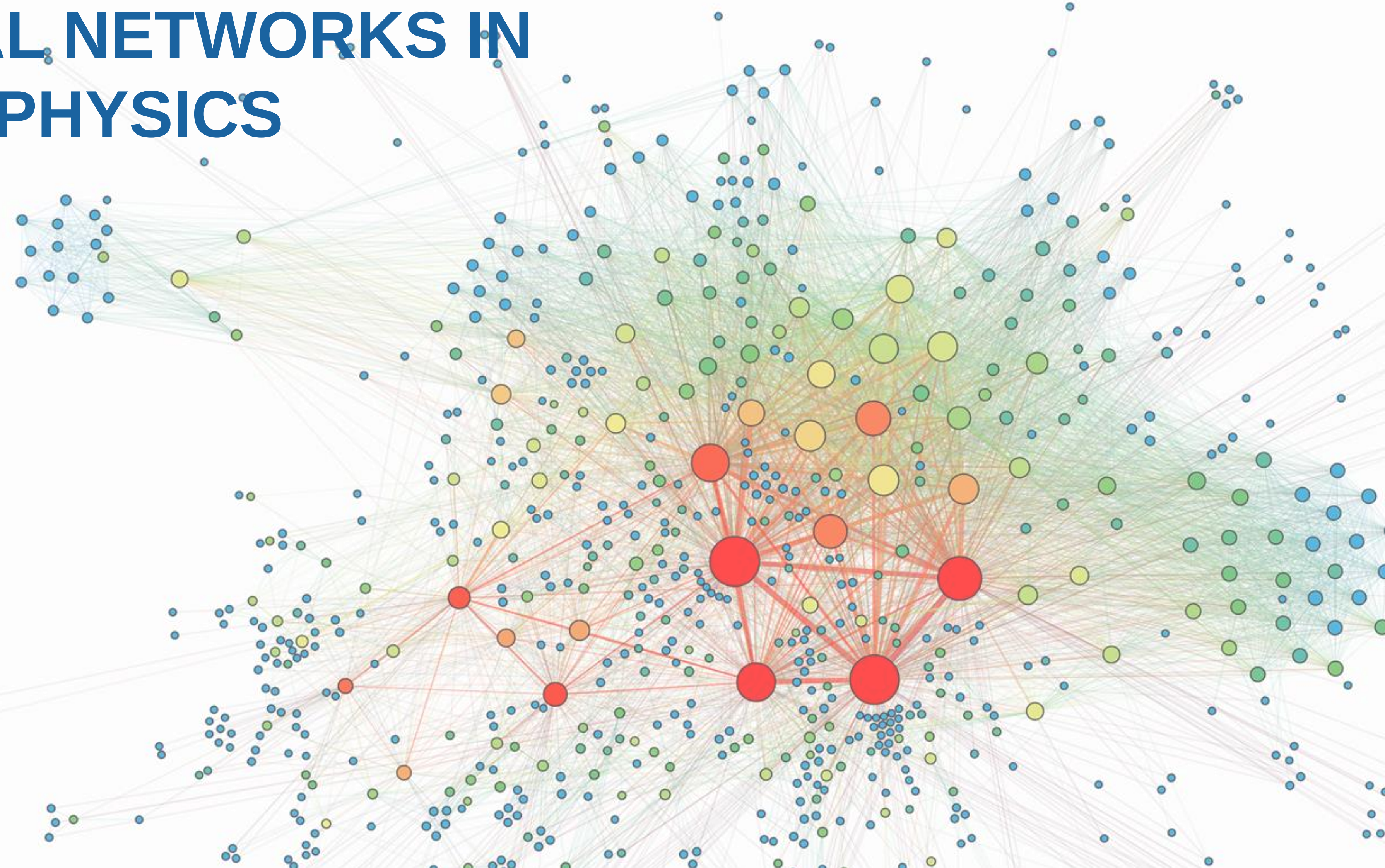
GRAPH DATA STRUCTURES AND GRAPH NEURAL NETWORKS IN HIGH ENERGY PHYSICS

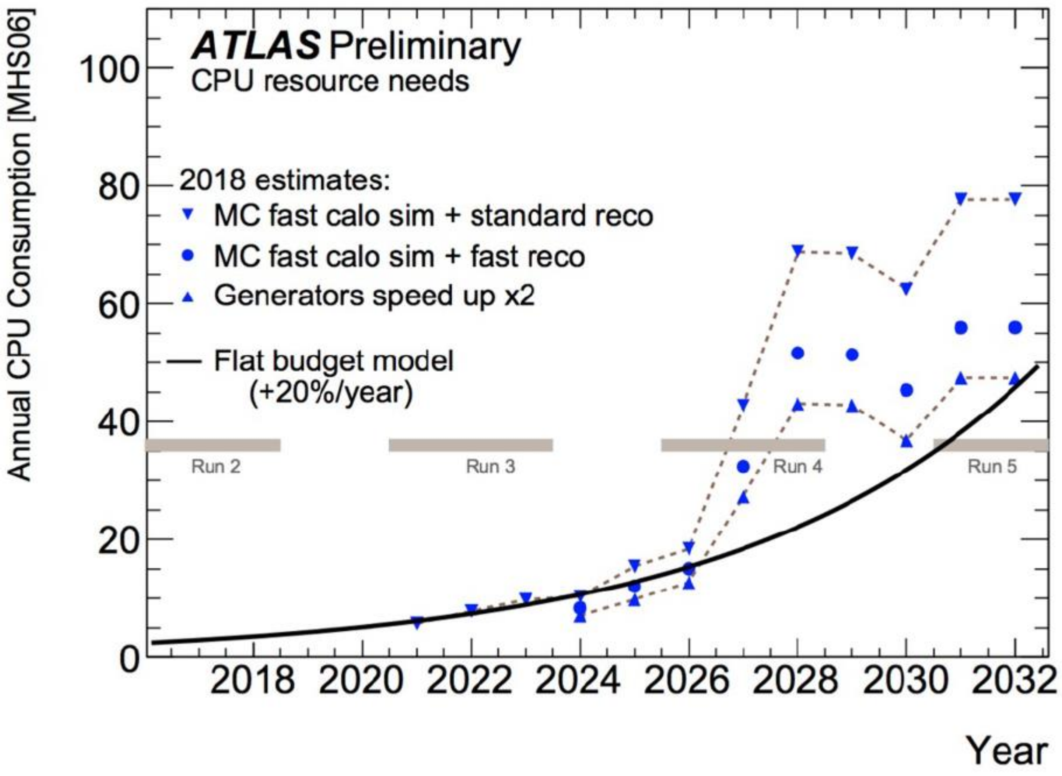
DANIEL MURNANE

ON BEHALF OF THE
EXA.TRKX COLLABORATION



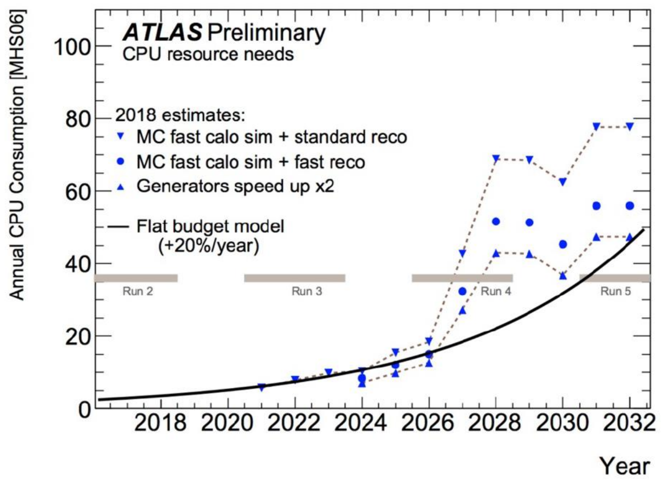
WHY GRAPHS?

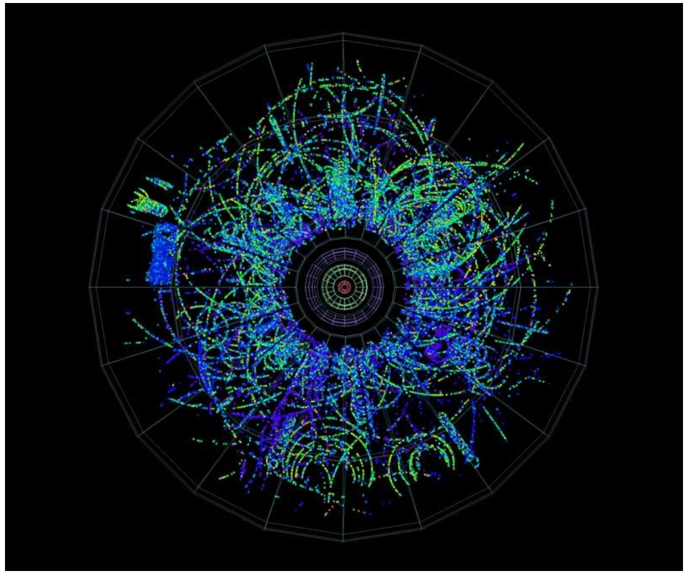
- High-luminosity scaling problem, means we need *something* complimenting traditional tracking algorithms,
  but why graphs?

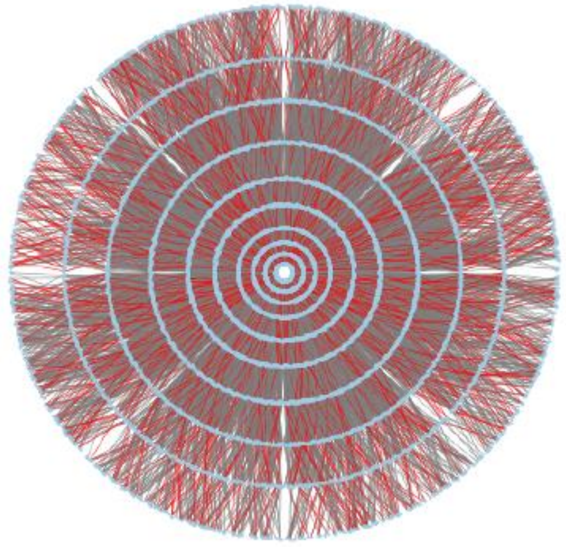


WHY GRAPHS?

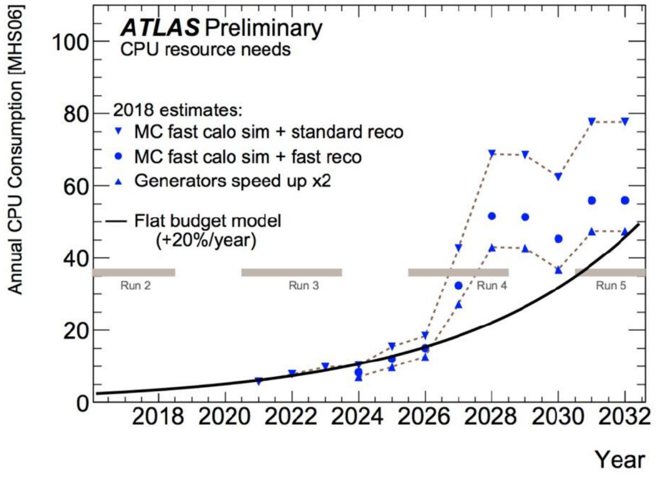
- High-luminosity scaling problem, means we need *something* complimenting traditional tracking algorithms, but why graphs?
- Graphs can capture inherent **sparsity** of much physics data







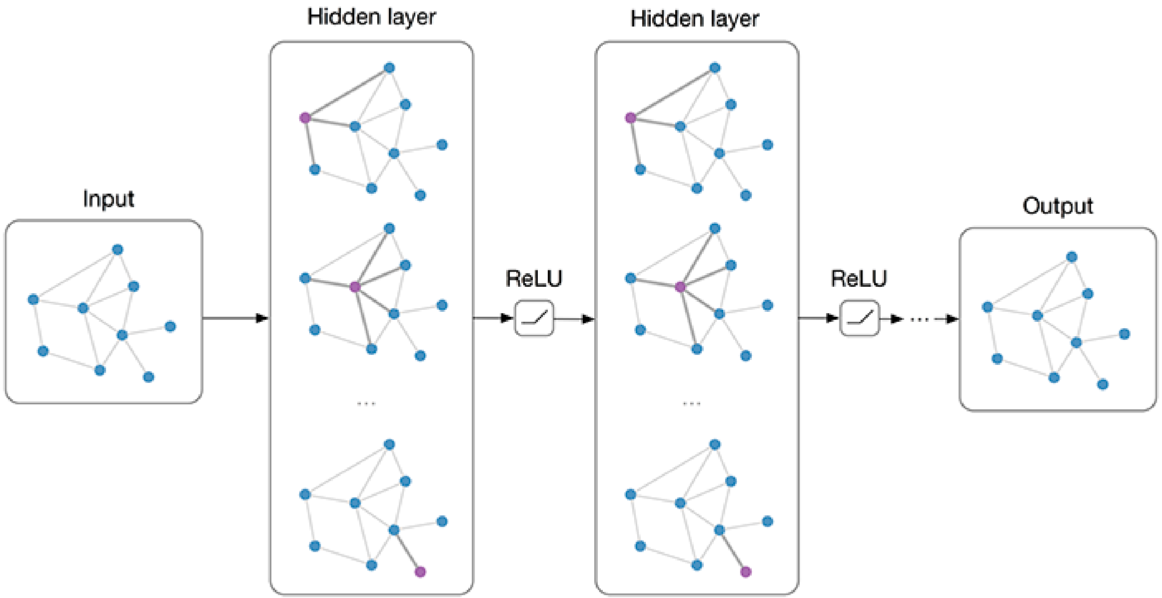
WHY GRAPHS?

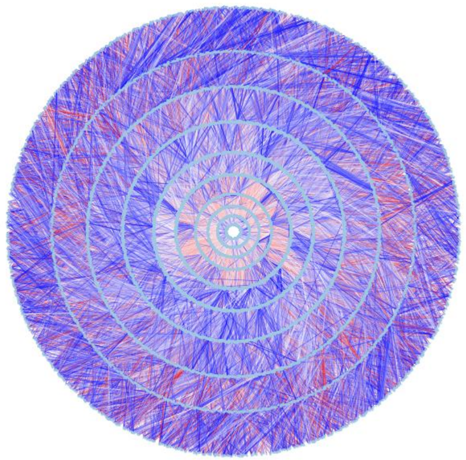


- High-luminosity scaling problem, means we need *something* complimenting traditional tracking algorithms, but why graphs?
- Graphs can capture inherent **sparsity** of much physics data
- Graphs can capture the **manifold** and **relational structure** of much physics data
- Conversion to and from graphs can allow **manipulation of dimensionality**
- **Graph Neural Networks** are booming (i.e. wouldn't be talking about graphs if there weren't powerful new methods to handle them)
- Industry **research and investment** means good outlook for software and hardware optimised for graphs

WHY GRAPH NEURAL NETWORKS?

- Can approximate geometry of the physics problem
- Are a generalisation of many other machine learning techniques
- E.g. Message passing convolution generalises CNN from flat to arbitrary geometry

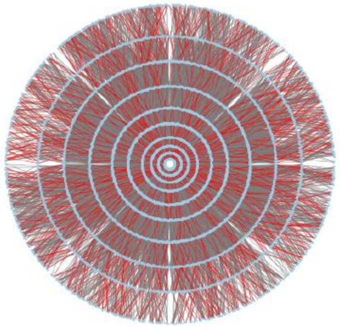




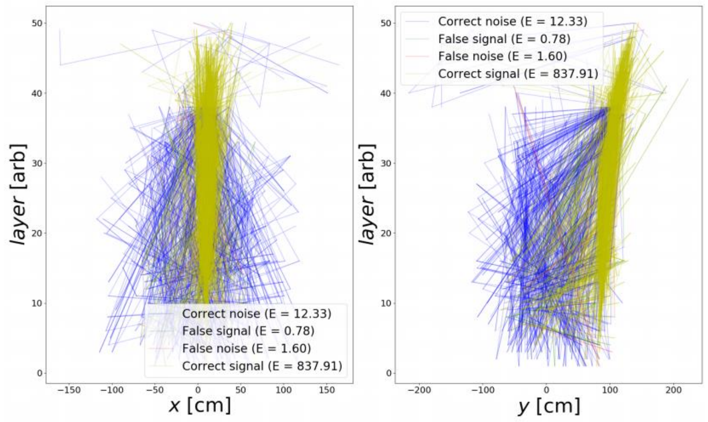
- Can learn node (i.e. hit / spacepoint) features and embeddings, as well as edge (i.e. relational) features and embeddings
- E.g. In practice, for a LHC-like detector environment, join hits into graph, and iterate through message-passing of hidden features

APPLICATIONS

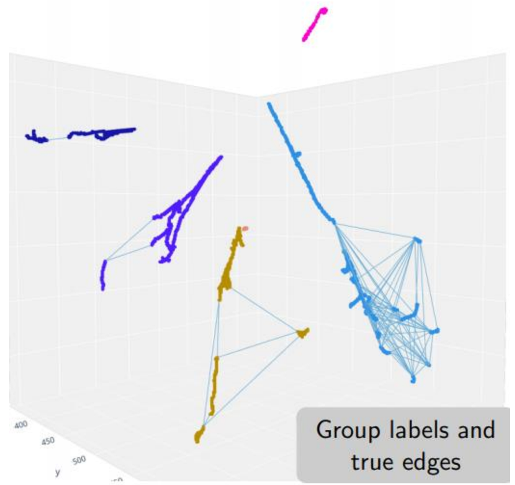
- TrackML dataset ~ HL-LHC silicon
  https://indico.cern.ch/event/831165/contributions/3717124/



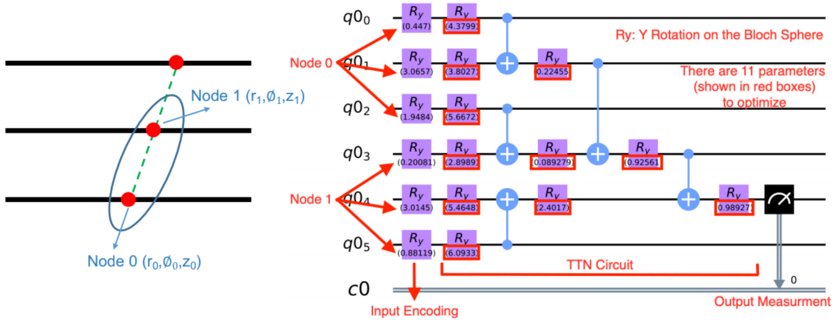
- High Granularity Calorimeter data
  https://arxiv.org/abs/2003.11603



- LArTPC data ~ DUNE experiment
  https://indico.cern.ch/event/831165/contributions/3717138/

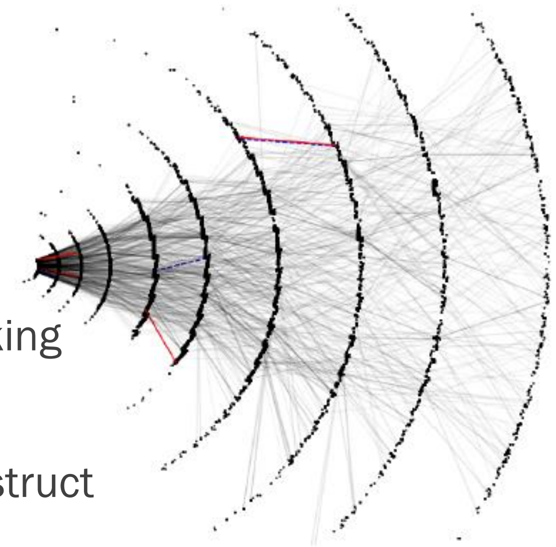


- Quantum GNN for Particle Track Reconstruction
  https://indico.cern.ch/event/831165/contributions/3717116/



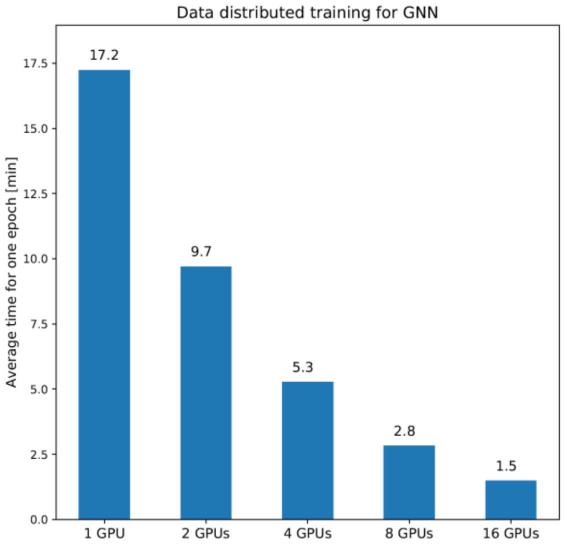
- GNNs on FPGAs for Level-1 Trigger
  https://indico.cern.ch/event/831165/contributions/3758961/

PERFORMANCE

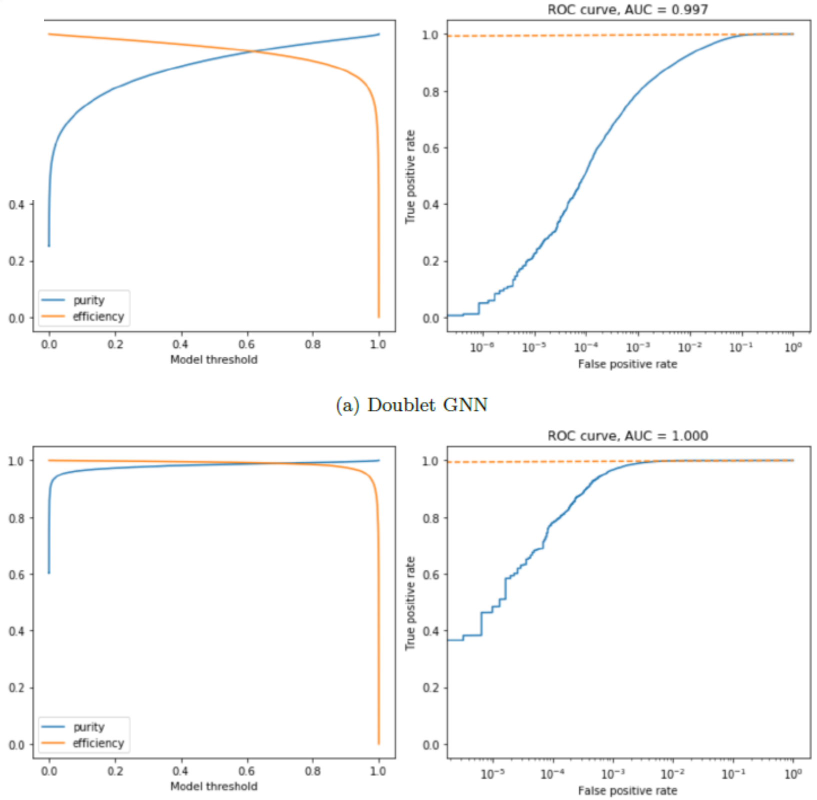


Example classified graph – correct (grey), incorrect (red)

- Accuracy metrics: Competitive with highest-ranking TrackML entries
- Timing metrics: average of 0.34 s/event to construct and classify (doublet) graph



- Distributed inference/training: scales as expected
- Scaling w/ luminosity: Less than quadratic with embedding-space construction and sparse message-passing operations



(a) Doublet GNN

(b) Triplet GNN

OUTLOOK

- Converging on better architectures (attention, gated RNN, generalising dense, flat methods to sparse, graph structure - not that the two are mutually inclusive, there is increasing interest in sparse CNN techniques, for example)

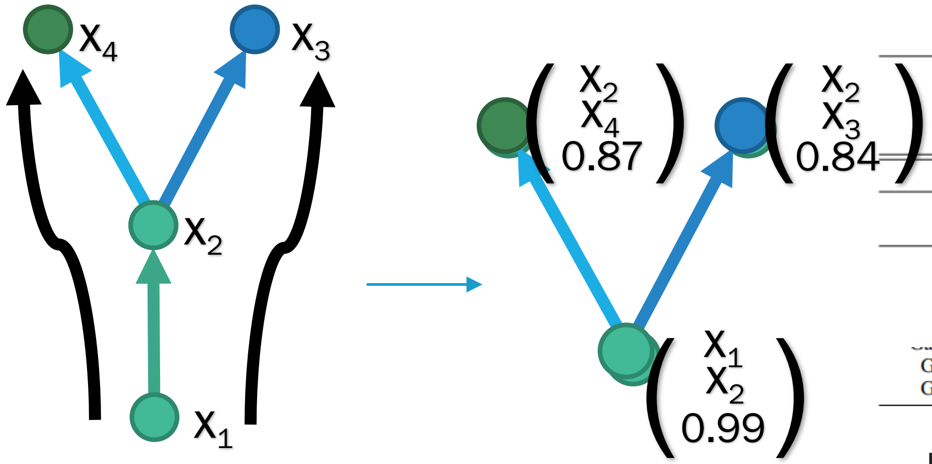
		LINK PREDICTION									
		TSP					COLLAB				
Model	L	**#Param**	**Test F1±s.d.**	**Train F1±s.d.**	**#Epoch**	**Epoch/Total**	**#Param ($L=3$)**	**Test Hits±s.d.**	**Train Hits±s.d.**	**#Epoch**	**Epoch/Total**
MLP	4	96956	0.544±0.001	0.544±0.001	164.25	50.15s/2.31hr	39441	20.350±2.168	29.807±3.360	147.50	2.09s/0.09hr
GCN	4	95702	0.630±0.001	0.631±0.001	261.00	152.89s/11.15hr	40479	50.422±1.131	92.112±0.991	122.50	351.05s/12.04hr
GraphSage	4	99263	0.665±0.003	0.669±0.003	266.00	157.26s/11.68hr	39856	51.618±0.690	99.949±0.052	152.75	277.93s/11.87hr
...											
GatedGCN-E	4	97858	0.808±0.003	0.811±0.003	197.00	218.51s/12.04hr	40965	49.212±1.560	88.747±1.058	95.00	451.21s/12.03hr
GatedGCN-E	16	500770	**0.838±0.002**	0.850±0.001	53.00	807.23s/12.17hr			-		

Dwivedi, Vijay Prakash, et al. "Benchmarking graph neural networks." *arXiv preprint arXiv:2003.00982* (2020).

OUTLOOK

- Converging on better architectures (attention, gated RNN, generalising dense, flat methods to sparse, graph structure - not that the two are mutually inclusive, there is increasing interest in sparse CNN techniques, for example)
- Converging on better methods (sparse operations, triplet graph structure, fast clustering, approximate NN, piggy-backing off big tech methods, e.g. Facebook FAISS)

Google Trends of “Graph Neural Networks”



Doublet graph to higher-order classification

			LINK PREDICTION									
			TSP					COLLAB				
Model	L	#Param	Test F1±s.d.	Train F1±s.d.	#Epoch	Epoch/Total	#Param ($L = 3$)	Test Hits±s.d.	Train Hits±s.d.	#Epoch	Epoch/Total
MLP	4	96956	0.544±0.001	0.544±0.001	164.25	50.15s/2.31hr	39441	20.350±2.168	29.807±3.360	147.50	2.09s/0.09hr
GCN	4	95702	0.630±0.001	0.631±0.001	261.00	152.89s/11.15hr	40479	50.422±1.131	92.112±0.991	122.50	351.05s/12.04hr
GraphSage	4	99263	0.665±0.003	0.669±0.003	266.00	157.26s/11.68hr	39856	51.618±0.690	99.949±0.052	152.75	277.93s/11.87hr
...											
GatedGCN-E	4	97858	0.808±0.003	0.811±0.003	197.00	218.51s/12.04hr	40965	49.212±1.560	88.747±1.058	95.00	451.21s/12.03hr
GatedGCN-E	16	500770	**0.838±0.002**	0.850±0.001	53.00	807.23s/12.17hr			-		

Dwivedi, Vijay Prakash, et al. "Benchmarking graph neural networks." *arXiv preprint arXiv:2003.00982* (2020).

OUTLOOK

Google Trends of "Graph Neural Networks"

- Converging on better architectures (attention, gated RNN, generalising dense, flat methods to sparse, graph structure - not that the two are mutually inclusive, there is increasing interest in sparse CNN techniques, for example)
- Converging on better methods (sparse operations, triplet graph structure, fast clustering, approximate NN, piggy-backing off big tech methods, e.g. Facebook FAISS)
- Converging on better hardware (mixed precision handling on new GPUs/TPUs, sparse handling in IPUs , compilability of graph-structure ML libraries for FPGA ports, e.g. IEEE HPEC GraphChallenge)

		LINK PREDICTION									
		TSP					COLLAB				
Model	L	**#Param**	**Test F1±s.d.**	**Train F1±s.d.**	**#Epoch**	**Epoch/Total**	**#Param ($L = 3$)**	**Test Hits±s.d.**	**Train Hits±s.d.**	**#Epoch**	**Epoch/Total**
MLP	4	96956	0.544±0.001	0.544±0.001	164.25	50.15s/2.31hr	39441	20.350±2.168	29.807±3.360	147.50	2.09s/0.09hr
GCN	4	95702	0.630±0.001	0.631±0.001	261.00	152.89s/11.15hr	40479	50.422±1.131	92.112±0.991	122.50	351.05s/12.04hr
GraphSage	4	99263	0.665±0.003	0.669±0.003	266.00	157.26s/11.68hr	39856	51.618±0.690	99.949±0.052	152.75	277.93s/11.87hr
...											
GatedGCN-E	4	97858	0.808±0.003	0.811±0.003	197.00	218.51s/12.04hr	40965	49.212±1.560	88.747±1.058	95.00	451.21s/12.03hr
GatedGCN-E	16	500770	**0.838±0.002**	0.850±0.001	53.00	807.23s/12.17hr			-		

Dwivedi, Vijay Prakash, et al. "Benchmarking graph neural networks." *arXiv preprint arXiv:2003.00982* (2020).